Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 98 (2007) 1180 – 1194
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Nonnegative quadratic estimation and quadratic sufficiency in general linear models

Xu-Qing Liu[a,*], Jian-Ying Rong[b]

[a]*Department of Computing Science, Huaiyin Institute of Technology, Huaian 223001, P.R. China*
[b]*Department of Foundation Courses, Huaian College of Information Technology, Huaian 223003, P.R. China*

Received 25 August 2005
Available online 26 September 2006

Abstract

Notions of linear sufficiency and quadratic sufficiency are of interest to some authors. In this paper, the problem of nonnegative quadratic estimation for $\beta'\mathbf{H}\beta + h\sigma^2$ is discussed in a general linear model and its transformed model. The notion of quadratic sufficiency is considered in the sense of generality, and the corresponding necessary and sufficient conditions for the transformation to be quadratically sufficient are investigated. As a direct consequence, the result on (ordinary) quadratic sufficiency is obtained. In addition, we pose a practical problem and extend a special situation to the multivariate case. Moreover, a simulated example is conducted, and applications to a model with compound symmetric covariance matrix are given. Finally, we derive a remark which indicates that our main results could be extended further to the quasi-normal case.

AMS 1991 subject classification: 62H12; 62J12; 15A04; 15A09

Keywords: Original model; Transformed model; Linear sufficiency; Quadratic sufficiency

1. Introduction

For convenience throughout this paper, we will write $\mathbf{A} \in \mathbb{R}_{m,n}$ if $\mathbf{A}$ is an $m \times n$ real matrix, $\mathbf{A} \in \mathbb{R}_n^{\mathrm{s}}$, if $\mathbf{A} \in \mathbb{R}_{n,n}$ and is symmetric, $\mathbf{A} \in \mathbb{R}_n^{\geqslant}$, if $\mathbf{A} \in \mathbb{R}_n^{\mathrm{s}}$ and is nonnegative definite. For given $\mathbf{A} \in \mathbb{R}_{m,n}$ denote by the symbols $\mathbf{A}'$, $\mathbf{A}^-$, $\mathbf{A}^+$, $\mathscr{R}(\mathbf{A})$, $\mathrm{rk}(\mathbf{A})$ and $P_{\mathbf{A}} = \mathbf{A}\mathbf{A}^+$ the transpose, any generalized inverse, the Moore–Penrose inverse, the column space (range), the rank and the orthogonal projection, respectively, of $\mathbf{A}$. In addition, $\mathbf{A}^{\perp}$ refers to a matrix of maximum rank

* Corresponding author.
E-mail addresses: seunuaalr_98@yahoo.com.cn (X. -Q. Liu), rjy98@163.com (J.-Y Rong).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.08.008

such that $\mathbf{A}'\mathbf{A}^{\perp} = \mathbf{0}$. Further, for given $\mathbf{A} \in \mathbb{R}_{n,n}$, let $\mathrm{tr}\,(\mathbf{A})$ be the trace of $\mathbf{A}$. Furthermore, for given $\mathbf{A}, \mathbf{B} \in \mathbb{R}_n^{\geqslant}$ the symbol $\mathbf{A} \leqslant \mathbf{B}$ (or $\mathbf{B} \geqslant \mathbf{A}$) will denote that $\mathbf{B} - \mathbf{A} \in \mathbb{R}_n^{\geqslant}$. Moreover, for any $\mathbf{A} \in \mathbb{R}_n^{s}$, $\lambda_i\,(\mathbf{A})$ denotes the ith largest eigenvalue of $\mathbf{A}$.

Consider a general linear model, denoted by

$$\mathscr{L} = \left\{ y,\ \mathbf{X}\beta,\ \sigma^2\mathbf{V} \right\}, \tag{1.1}$$

where y is an n-dimensional normally distributed random vector of observations, with the expectation vector $\mathscr{E}\,(y) = \mathbf{X}\beta$ and covariance matrix $\mathscr{D}\,(y) = \sigma^2\mathbf{V}$, where the nonstochastic matrix $\mathbf{X} \in \mathbb{R}_{n,p}$ with rank r $(\leqslant p < n)$ and $\mathbf{V} \in \mathbb{R}_n^{\geqslant}$ are supposed to be known, while $\beta \in \mathbb{R}_{p,1}$ and $\sigma^2 \in (0, +\infty)$ are unknown parameters. In some situations, the observation vector y may not be available, however, the vector $\mathbf{F}y$ is obtainable, where $\mathbf{F} \in \mathbb{R}_{m,n}$ is a known matrix with rank $\tilde{s}$ and $\mathrm{rk}\,(\mathbf{FX}) = \tilde{r}$, $m \geqslant p$. Thus, we have the following transformed model:

$$\mathscr{L}_T = \left\{ \mathbf{F}y,\ \mathbf{FX}\beta,\ \sigma^2\mathbf{FVF}' \right\}. \tag{1.2}$$

If the interest is in estimating $\mathbf{X}\beta$ it is reasonable to consider the so-called linearly sufficient estimation as defined in Drygas [3], see also among others. The notion of linear sufficiency introduced by Drygas [3] is that $\mathbf{F}y$ is said linearly sufficient if there is a linear function of $\mathbf{F}y$ which is the best linear unbiased (BLU) estimator of $\mathbf{X}\beta$. This notion has been considered by many statisticians, among them Drygas [4,5], Baksalary and Mathew [2], Müeller [16], Heiligers and Markiewicz [11], Markiewicz [15] are mentioned. And for a more generalization one can see Ip et al. [13]. Another similar concept is the so-called quadratic sufficiency; e.g., cf. [1].

In the past several decades, some authors have considered the problem of the nonnegative estimation for a quadratic parametric function, namely $\theta = \beta'\mathbf{H}\beta + h\sigma^2$ with $\mathbf{H} \in \mathbb{R}_p^{\geqslant}$ and $h \geqslant 0$. The problem arises, for instance, if one wants to estimate the accuracy of linear estimators of $\mathbf{X}\beta$ by means of the mean squared error (MSE) or to predict the mean of $y'y$, which have the structures of θ.

Here, we will consider a notion of quadratic sufficiency of generality. The left of the paper is as follows. In Section 2, the notion is discussed in two cases and the corresponding necessary and sufficient conditions for the transformation to be quadratically sufficient are derived. Further, some special results are given, in which the (ordinary) quadratic sufficiency is obtained as a direct consequence. In Section 3, a problem in practice and the multivariate case are mentioned. In addition, we give a simulated example, which in some sense argues for our conclusions. Then, in Section 4, we apply our main results to a linear model with covariance matrix being compound symmetric. Moreover, we demonstrate an intuitional impression that transformation matrix, $\mathbf{F}$, being full column rank is sufficient (truly) and yet necessary for $\mathscr{L}_T$ to contain all information of $\mathscr{L}$. Finally, we give a remark which indicates that our main results could be extended further to the quasi-normal case in Section 5.

2. Main results on quadratic sufficiency

First of all, we introduce some notations as follows. Let

$$\mathbf{W} = \mathbf{V} + \mathbf{XUX}', \quad \widetilde{\mathbf{W}} = \mathbf{FWF}',$$

with $\mathrm{rk}\,(\mathbf{W}) = s$ and $\mathbf{U} \in \mathbb{R}_p^{\geqslant}$ such that $\mathscr{R}\,(\mathbf{W}) = \mathscr{R}\,(\mathbf{X}, \mathbf{V})$, or equivalently, $\mathscr{R}\,(\mathbf{X}) \subseteq \mathscr{R}\,(\mathbf{W})$, which further implies $\mathscr{R}\left(\widetilde{\mathbf{W}}\right) = \mathscr{R}\left(\mathbf{FX}, \mathbf{FVF}'\right)$. In the following we will assume $\mathscr{R}\left(\mathbf{F}'\right) \subseteq \mathscr{R}\,(\mathbf{W})$

as argued in Müeller [16] and Ip et al. [13] that one can neglect all elements outside $\mathscr{R}(\mathbf{W})$ since they form a null set of the linear model and our statistical inference is based on the linear statistic $\mathbf{F}y$.

It should be noticed that $\mathscr{R}(\mathbf{F}) = \mathscr{R}(\widetilde{\mathbf{W}})$ in terms of the implication

$$\mathscr{R}(\mathbf{F}') \subseteq \mathscr{R}(\mathbf{W}) = \mathscr{R}\left(\mathbf{W}^{\frac{1}{2}}\right) \Rightarrow \mathscr{R}(\mathbf{F}) = \mathscr{R}(\mathbf{F}\mathbf{F}') \subseteq \mathscr{R}\left(\mathbf{F}\mathbf{W}^{\frac{1}{2}}\right) = \mathscr{R}(\widetilde{\mathbf{W}}),$$

since $\mathscr{R}(\mathbf{W}) \subseteq \mathscr{R}(\widetilde{\mathbf{F}})$ holds inherently, and therefore $\mathrm{rk}(\widetilde{\mathbf{W}}) = \mathrm{rk}(\mathbf{F}) = \tilde{s}$. Without loss of generality, we assume $s > r$ and $\tilde{s} > \tilde{r}$. Following Rao's *Unified Theory of Least Squares* we now denote by $\mathbf{X}\hat{\beta}$ and $\hat{\sigma}^2$ the unbiased estimators, respectively, of $\mathbf{X}\beta$ and σ^2 in the original model, while by $\mathbf{F}\mathbf{X}\tilde{\beta}$ and $\tilde{\sigma}^2$ the unbiased estimators, respectively, of $\mathbf{F}\mathbf{X}\beta$ and σ^2 in the transformed model, with

$$\begin{aligned}
\hat{\beta} &= \left(\mathbf{X}'\mathbf{W}^-\mathbf{X}\right)^- \mathbf{X}'\mathbf{W}^- y,\\
\tilde{\beta} &= \left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^-\mathbf{F}\mathbf{X}\right)^- \mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^-\mathbf{F}y,\\
\hat{\sigma}^2 &= \left(y - \mathbf{X}\hat{\beta}\right)' \mathbf{W}^- \left(y - \mathbf{X}\hat{\beta}\right) \Big/ (s - r),\\
\tilde{\sigma}^2 &= \left(y - \mathbf{X}\tilde{\beta}\right)' \mathbf{F}'\widetilde{\mathbf{W}}^-\mathbf{F} \left(y - \mathbf{X}\tilde{\beta}\right) \Big/ (\tilde{s} - \tilde{r}).
\end{aligned}$$

Here it would be noticed that $\mathbf{X}\hat{\beta}$, $\mathbf{F}\mathbf{X}\tilde{\beta}$, $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ are invariant with probability one to the choices of the involved generalized inverses since $y \in \mathscr{R}(\mathbf{W})$ holds almost surely. Without loss of generality, we can replace the involved generalized inverses by the corresponding Moore–Penrose inverses in the above four representations whenever it is necessary, and vice versa.

Thus, for the nonnegative quadratic estimable function $\beta'\mathbf{H}\beta + h\sigma^2$ (in the sense that there is some quadratic form of y being unbiased for $\beta'\mathbf{H}\beta + h\sigma^2$, e.g., cf. [6], that is to see, $P_{\mathbf{X}'}\mathbf{H}P_{\mathbf{X}'} = \mathbf{H}$ under model $\mathscr{L}$ and $P_{\mathbf{X}'\mathbf{F}'}\mathbf{H}P_{\mathbf{X}'\mathbf{F}'} = \mathbf{H}$ under model $\mathscr{L}_T$), we have two naive estimators $\hat{\theta}_1 = \hat{\beta}'\mathbf{H}\hat{\beta} + h\hat{\sigma}^2$ and $\tilde{\theta}_1 = \tilde{\beta}'\mathbf{H}\tilde{\beta} + h\tilde{\sigma}^2$ for θ. Note that it is appealing, that the condition $P_{\mathbf{X}'}\mathbf{H}P_{\mathbf{X}'} = \mathbf{H}$ is equivalent to $P_{\mathbf{X}'}\mathbf{H} = \mathbf{H}P_{\mathbf{X}'} = \mathbf{H}$ while $P_{\mathbf{X}'\mathbf{F}'}\mathbf{H}P_{\mathbf{X}'\mathbf{F}'} = \mathbf{H} \Leftrightarrow P_{\mathbf{X}'\mathbf{F}'}\mathbf{H} = \mathbf{H}$. The following lemma concerns the above two estimators.

Lemma 2.1. *Assume θ is nonnegative quadratic estimable under $\mathscr{L}$ and $\mathscr{L}_T$. Then $\hat{\theta}_1$ and $\tilde{\theta}_1$ are potentially biased for θ with*

$$\begin{aligned}
\mathrm{Bias}\left(\hat{\theta}_1, \theta\right) &= \sigma^2 \,\mathrm{tr}\left\{\mathbf{H}\left[\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ - \mathbf{U}\right]\right\} \geqslant 0,\\
\mathrm{Bias}\left(\tilde{\theta}_1, \theta\right) &= \sigma^2 \,\mathrm{tr}\left\{\mathbf{H}\left[\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^+\mathbf{F}\mathbf{X}\right)^+ - \mathbf{U}\right]\right\} \geqslant 0.
\end{aligned}$$

Proof. For convenience, we write $\left(\mathbf{A}^+\right)^{\frac{1}{2}}$ as $\mathbf{A}^{+\frac{1}{2}}$ for any nonnegative definite matrix $\mathbf{A}$, the same below. Firstly, it is clear that $\mathscr{E}\left(\hat{\sigma}^2\right) = \sigma^2$ since $\hat{\sigma}^2$ is unbiased for σ^2. Note that $\mathscr{R}(\mathbf{X}') = \mathscr{R}(\mathbf{X}'\mathbf{W}^+\mathbf{X})$ in terms of the implication

$$\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}\left(\mathbf{W}^{+\frac{1}{2}}\right) \Rightarrow \mathscr{R}(\mathbf{X}') = \mathscr{R}(\mathbf{X}'\mathbf{X}) \subseteq \mathscr{R}\left(\mathbf{X}'\mathbf{W}^{+\frac{1}{2}}\right) = \mathscr{R}(\mathbf{X}'\mathbf{W}^+\mathbf{X}),$$

since $\mathscr{R}(\mathbf{X}'\mathbf{W}^+\mathbf{X}) \subseteq \mathscr{R}(\mathbf{X}')$ holds inherently. Further, $P_{\mathbf{X}'} = P_{\mathbf{X}'\mathbf{W}^+\mathbf{X}}$. Employing the fact that if η is some random vector with expectation vector μ and covariance matrix $\mathbf{\Sigma}$, then

$$\mathscr{E}\left(\eta'\mathbf{D}\eta\right) = \mu'\mathbf{D}\mu + \mathrm{tr}(\mathbf{D}\mathbf{\Sigma}) \tag{2.1}$$

for some given symmetric matrix $\mathbf{D}$ of suitable order, and recalling that $P_{\mathbf{X}'}\mathbf{H} = \mathbf{H}$, one can conclude that $\mathscr{E}\left(\hat{\theta}_1\right) = \theta + \sigma^2 \operatorname{tr}\{\mathbf{H}[(\mathbf{X}'\mathbf{W}^+\mathbf{X})^+ - \mathbf{U}]\}$ by direct operations and therefore

$$\operatorname{Bias}\left(\hat{\theta}_1, \theta\right) = \sigma^2 \operatorname{tr}\left\{\mathbf{H}^{\frac{1}{2}}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ \mathbf{X}'\mathbf{W}^+\mathbf{V}\mathbf{W}^+\mathbf{X}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ \mathbf{H}^{\frac{1}{2}}\right\} \geqslant 0,$$

which completes the proving since the latter follows directly from the former. □

Note that the estimators $\hat{\theta} = \hat{\theta}_1 - \operatorname{Bias}\left(\hat{\theta}_1, \theta\right)\hat{\sigma}^2 \Big/ \sigma^2$ and $\tilde{\theta} = \tilde{\theta}_1 - \operatorname{Bias}\left(\tilde{\theta}_1, \theta\right)\tilde{\sigma}^2 \Big/ \sigma^2$ are unbiased for θ, by Lemma 2.1. And what's more, one can conclude that $\hat{\theta}$ (resp., $\tilde{\theta}$) is the essentially unique best quadratic unbiased (BQU) estimator under $\mathscr{L}$ (resp., $\mathscr{L}_T$); cf. Appendix A.

By Lemma 2.1, we notice that the values of $\hat{h} = h - \operatorname{tr}\{\mathbf{H}[(\mathbf{X}'\mathbf{W}^+\mathbf{X})^+ - \mathbf{U}]\}$ and $\tilde{h} = h - \operatorname{tr}\{\mathbf{H}[(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^+\mathbf{F}\mathbf{X})^+ - \mathbf{U}]\}$ can be either negative or nonnegative. Indeed, if $\hat{h}$ is nonnegative, then we are readily adopting the estimator $\hat{\theta}$, however, it is unacceptable in practice if $\hat{h}$ is negative since the value of $\hat{\theta}$ can be negative. Then we would like to use the so-called nonnegative minimum-biased (NNMB) estimation or called nonnegative best quadratic-biased (NN-BQB) estimation for θ. $\tilde{\theta}$ can be discussed with similar fashion.

Let us now consider the notion of quadratic sufficiency (in the sense for any point $\left(\beta, \sigma^2\right) \in \mathbb{R}_{p,1} \times (0, +\infty)$) of generality. The following definition concerns when model $\mathscr{L}_T$ contains all information of model $\mathscr{L}$ in some sense.

Definition 2.1. Assume $\mathscr{C}$ is a given class of nonnegative quadratic estimation for θ. $\mathbf{F}y$ is said quadratically sufficient (with respect to the class $\mathscr{C}$) if there exists a symmetric matrix $\mathbf{G}$ such that $y'\mathbf{F}'\mathbf{G}\mathbf{F}y$ is the best estimator for θ in the class $\mathscr{C}$. We write $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C})}$ if $\mathbf{F}y$ is quadratically sufficient with respect to $\mathscr{C}$.

2.1. Nonnegative quadratic unbiased estimation and quadratic sufficiency

As one can see, according to Definition 2.1, taking $\mathscr{C} = \mathscr{C}_0$ for given $\mathscr{C}_0$ gives corresponding quadratic sufficiency. We first consider this notion in the case $\mathscr{C} = \mathscr{C}_1$, namely the class of nonnegative quadratic unbiased estimation for θ, when both $\hat{h}$ and $\tilde{h}$ are nonnegative in this subsection. Note that the implication $\tilde{h} \geqslant 0 \Rightarrow \hat{h} \geqslant 0$ holds inherently provided θ is nonnegative quadratic estimable under model $\mathscr{L}_T$ (and thereby under model $\mathscr{L}$); see also the proof of Theorem 2.2.

It should be noticed that inserting $\mathbf{H} = \mathbf{0}$ and $h = s - r$ into Definition 2.1 gives the notion as defined in Baksalary and Drygas [1] that $\mathbf{F}y$ is quadratically sufficient if there exists a symmetric matrix $\mathbf{G}$ such that $y'\mathbf{F}'\mathbf{G}\mathbf{F}y$ is the BQU estimator for $(s - r)\,\sigma^2$, which coincides with the so-called linear error-sufficiency; cf. [9].

According to Appendix A , Definition 2.1 with $\mathscr{C} = \mathscr{C}_1$ is equivalent to that $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ if and only if $\hat{\theta}$ and $\tilde{\theta}$ are identical almost surely since the BQU estimator is essentially unique. Before giving the necessary and sufficient conditions for $\hat{\theta}$ and $\tilde{\theta}$ to be identical, we first deduce the covariance between $\hat{\theta}$ and $\tilde{\theta}$. Denote

$$\hat{b} = \left[h - \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ - \mathbf{H}\mathbf{U}\right\}\right] \Big/ (s - r),$$

$$\widehat{\mathbf{A}} = \mathbf{W}^+\mathbf{X}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ \mathbf{H}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ \mathbf{X}'\mathbf{W}^+,$$

$$\widehat{\mathbf{B}} = \mathbf{W}^+ - \mathbf{W}^+\mathbf{X}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ \mathbf{X}'\mathbf{W}^+$$

and

$$\tilde{b} = \left[h - \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+} - \mathbf{H}\mathbf{U}\right\}\right] \Big/ (\tilde{s} - \tilde{r}),$$
$$\widetilde{\mathbf{A}} = \mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F},$$
$$\widetilde{\mathbf{B}} = \mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F} - \mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}.$$

Employing the fact that if $\eta \sim \mathcal{N}_n(\mu, \mathbf{\Sigma})$ then

$$\operatorname{Cov}\left(\eta'\mathbf{A}\eta, \eta'\mathbf{B}\eta\right) = 2\operatorname{tr}(\mathbf{A\Sigma B\Sigma}) + 4\mu'\mathbf{A\Sigma B}\mu \tag{2.2}$$

for given nonstochastic matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}_n^s$, which can be deduced following from the fact that if $\mathbf{\Sigma} \geqslant 0$ with rank t, then $\eta \sim \mathcal{N}_n(\mu, \mathbf{\Sigma})$ if and only if η has stochastic decomposition $\eta = \mu + \mathbf{C}\xi$ with $\xi \sim \mathcal{N}_t(\mathbf{0}, \mathbf{I}_t)$ and $\mathbf{C} \in \mathbb{R}_{n,t}$ satisfying $\mathbf{\Sigma} = \mathbf{CC}'$, noting that $\widehat{\mathbf{B}}\mathbf{X} = \widetilde{\mathbf{B}}\mathbf{X} = \mathbf{0}$ and $\mathbf{X}'\widehat{\mathbf{A}}\mathbf{X} = \mathbf{X}'\widetilde{\mathbf{A}}\mathbf{X} = \mathbf{H}$, one can obtain

(1) $\operatorname{tr}\left(\widehat{\mathbf{A}}\mathbf{V}\widetilde{\mathbf{A}}\mathbf{V}\right) = \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+} - 2\mathbf{HUH}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+} + \mathbf{HUHU}\right\}$,
(2) $\operatorname{tr}\left(\widehat{\mathbf{A}}\mathbf{V}\widetilde{\mathbf{B}}\mathbf{V}\right) = \mathbf{0}$,
(3) $\operatorname{tr}\left(\widehat{\mathbf{B}}\mathbf{V}\widetilde{\mathbf{A}}\mathbf{V}\right) = \operatorname{tr}\left\{\mathbf{H}\left[\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+} - \left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\right]\right\}$,
(4) $\operatorname{tr}\left(\widehat{\mathbf{B}}\mathbf{V}\widetilde{\mathbf{B}}\mathbf{V}\right) = \tilde{s} - \tilde{r}$,
(5) $\mathbf{X}'\widehat{\mathbf{A}}\mathbf{V}\widetilde{\mathbf{A}}\mathbf{X} = \mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H} - \mathbf{HUH}$

by direct operations. It follows that

$$\begin{aligned}
\operatorname{Cov}\left(\hat{\theta}, \tilde{\theta}\right) &= \operatorname{Cov}\left[y'\left(\widehat{\mathbf{A}} + \hat{b}\widehat{\mathbf{B}}\right)y,\, y'\left(\widetilde{\mathbf{A}} + \tilde{b}\widetilde{\mathbf{B}}\right)y\right] \\
&= 2\sigma^4 \operatorname{tr}\left(\widehat{\mathbf{A}}\mathbf{V}\widetilde{\mathbf{A}}\mathbf{V}\right) + 2\tilde{b}\sigma^4 \operatorname{tr}\left(\widehat{\mathbf{A}}\mathbf{V}\widetilde{\mathbf{B}}\mathbf{V}\right) + 2\hat{b}\sigma^4 \operatorname{tr}\left(\widehat{\mathbf{B}}\mathbf{V}\widetilde{\mathbf{A}}\mathbf{V}\right) \\
&\quad + 2\hat{b}\tilde{b}\sigma^4 \operatorname{tr}\left(\widehat{\mathbf{B}}\mathbf{V}\widetilde{\mathbf{B}}\mathbf{V}\right) + 4\sigma^2\beta'\mathbf{X}'\widehat{\mathbf{A}}\mathbf{V}\widetilde{\mathbf{A}}\mathbf{X}\beta \\
&= 2\sigma^4 \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+} - 2\mathbf{HUH}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+} + \mathbf{HUHU}\right\} \\
&\quad + 2\hat{b}\sigma^4 \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+} - \mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\right\} \\
&\quad + 2\hat{b}\tilde{b}\sigma^4(\tilde{s} - \tilde{r}) + 4\sigma^2\beta'\mathbf{H}\left[\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+} - \mathbf{U}\right]\mathbf{H}\beta.
\end{aligned}$$

We observe that inserting $\mathbf{F} = \mathbf{I}_n$ into the expression of $\operatorname{Cov}(\hat{\theta}, \tilde{\theta})$ gives $\mathscr{D}(\hat{\theta})$ and further replacing $\mathbf{X}$ by $\mathbf{FX}$ in the presentation of $\mathscr{D}(\hat{\theta})$ yields $\mathscr{D}(\tilde{\theta})$. Specially, by direct operations one can readily verify that $\operatorname{Cov}(\hat{\theta}, \tilde{\theta}) = \mathscr{D}(\hat{\theta})$. The following theorems give some results concerning $\mathscr{S}_{(\mathscr{C}_1)}$.

Theorem 2.1. *Assume θ is nonnegative quadratic estimable under model $\mathscr{L}_T$. Then $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ if and only if the following conditions hold simultaneously*:

(a) $\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{H} = \mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H}$,
(b) $\hat{h}(s - \tilde{s} + \tilde{r} - r) = 0$.

Specially, when $\hat{h}$ is positive then (b) *reduces to $\tilde{s} - \tilde{r} = s - r$.*

Proof. It suffices to prove that $\hat{\theta}$ and $\tilde{\theta}$ are identical if and only if conditions (a) and (b) hold simultaneously. Actually, $\hat{\theta}$ coincides with $\tilde{\theta}$ if and only if $\mathscr{D}(\hat{\theta}) = \mathscr{D}(\tilde{\theta})$ for any point $\left(\beta, \sigma^2\right) \in \mathbb{R}_{p,1} \times (0, +\infty)$, since both $\hat{\theta}$ and $\tilde{\theta}$ are unbiased for θ and $\operatorname{Cov}(\hat{\theta}, \tilde{\theta}) = \mathscr{D}(\hat{\theta})$, which is further equivalent to the following conditions are satisfied simultaneously combining the algebraic fact that if $\mathbf{A}, \mathbf{B} \in \mathbb{R}_n^s$ and $x'\mathbf{A}x = x'\mathbf{B}x$ holds for any $x \in \mathbb{R}_{n,1}$ then $\mathbf{A} = \mathbf{B}$:

- $\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{H} = \mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H}$,
- $\hat{\varrho} = \tilde{\varrho}$, with notations

$$\hat{\varrho} = \operatorname{tr}\left[\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\right]^{2} - 2\operatorname{tr}\left[\mathbf{HUH}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\right] + \hat{b}^{2}(s-r),$$
$$\tilde{\varrho} = \operatorname{tr}\left[\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\right]^{2} - 2\operatorname{tr}\left[\mathbf{HUH}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\right] + \tilde{b}^{2}(\tilde{s}-\tilde{r}).$$

As we can see, the first presentation refers to (a), which combining the second expression gives $\tilde{h} = \hat{h}$ and therefore yields (b), and vice versa. Thus we complete the proving. □

The following theorem offers an alternative version to Theorem 2.1.

Theorem 2.2. *Assume θ is nonnegative quadratic estimable under model $\mathscr{L}_T$. Then $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ if and only if* (*b*), *given in Theorem* 2.1, *and* (a*)/(a$_*$), *stated below*, *are satisfied simultaneously*:

(a*) $\operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{H}\right\} = \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H}\right\}$,

(a$_*$) $\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H} = \mathbf{H}$.

Proof. Let us now first show that

$$P_{\mathbf{X}'\mathbf{F}'}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}P_{\mathbf{X}'\mathbf{F}'} \leqslant \left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}. \tag{2.3}$$

Actually, employing the fact (e.g., [19]) that if $\mathbf{A} \geqslant \mathbf{0}$ and $\mathbf{B} \geqslant \mathbf{0}$ then

$$\mathbf{A} \geqslant \mathbf{B} \quad \Leftrightarrow \quad \begin{cases} \mathscr{R}(\mathbf{B}) \subseteq \mathscr{R}(\mathbf{A}), \\ \mathbf{B} \geqslant \mathbf{B}\mathbf{A}^{+}\mathbf{B}, \end{cases}$$

writing $\mathbf{S} = \mathbf{X}'\mathbf{W}^{+}\mathbf{X}$ and $\widetilde{\mathbf{S}} = \mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}$, it follows that $\widetilde{\mathbf{S}} = \mathbf{X}'\mathbf{W}^{+\frac{1}{2}}P_{\mathbf{W}^{\frac{1}{2}}\mathbf{F}'}\mathbf{W}^{+\frac{1}{2}}\mathbf{X} \leqslant \mathbf{S}$ and therefore

$$P_{\mathbf{X}'\mathbf{F}'}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}P_{\mathbf{X}'\mathbf{F}'} - \left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+} = \widetilde{\mathbf{S}}^{+}\left(\widetilde{\mathbf{S}}\mathbf{S}^{+}\widetilde{\mathbf{S}} - \widetilde{\mathbf{S}}\right)\widetilde{\mathbf{S}}^{+} \leqslant \mathbf{0}.$$

That is to say (2.3) holds. Further,

$$\mathbf{H}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}\mathbf{H} = \mathbf{H}P_{\mathbf{X}'\mathbf{F}'}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}P_{\mathbf{X}'\mathbf{F}'}\mathbf{H} \leqslant \mathbf{H}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X}\right)^{+}\mathbf{H},$$

which combining the fact that if $\mathbf{B} \geqslant \mathbf{A} \geqslant \mathbf{0}$, then $\mathbf{A} = \mathbf{B} \Leftrightarrow \operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{B})$ yields (a*) from (a) and vice versa; The equivalence between (a$_*$) and (a) follows directly from the fact that if $\mathbf{B} \geqslant \mathbf{A} \geqslant \mathbf{0}$ then $\mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{X}'\mathbf{B}\mathbf{X} \Leftrightarrow \mathbf{A}\mathbf{X} = \mathbf{B}\mathbf{X}$. □

If the interest is in estimating σ^2, which is nonnegative quadratic estimable under $\mathscr{L}$ and $\mathscr{L}_T$ inherently, then the following conclusion is clear, which coincides with Theorem 2 in Groß [9], combining the facts that

$$\mathscr{R}\left[\mathbf{F}'(\mathbf{F}\mathbf{X})^{\perp}\right] = \mathscr{R}\left(\mathbf{F}'\right) \cap \mathscr{R}(\mathbf{X})^{\perp}, \quad \operatorname{rk}(\mathbf{A}, \mathbf{B}) = \operatorname{rk}(\mathbf{A}) + \operatorname{rk}\left[\left(\mathbf{I} - \mathbf{A}\mathbf{A}^{+}\right)\mathbf{B}\right].$$

Corollary 2.1. $\mathbf{F}y$ *is quadratically sufficient to the class of nonnegative quadratic unbiased estimation of* σ^2 *iff* $\tilde{s} - \tilde{r} = s - r$, *which is further equivalent to*

$$\mathscr{R}\left(\mathbf{V}\mathbf{X}^{\perp}\right) = \mathscr{R}\left[\mathbf{V}\mathbf{F}'\left(\mathbf{F}\mathbf{X}\right)^{\perp}\right].$$

Following the above theorems, let us now investigate some special and important situations, in which the case $\mathscr{R}\left(\mathbf{F}'\right) = \mathscr{R}\left(\mathbf{W}\right)$ is noticeable, not only mathematically interesting but also practically relevant. By means of the equivalence (cf. [19])

$$\mathbf{A} \geqslant \mathbf{B} \quad \Leftrightarrow \quad \begin{cases} \mathscr{R}\left(\mathbf{B}\right) \subseteq \mathscr{R}\left(\mathbf{A}\right), \\ \lambda_1\left(\mathbf{B}\mathbf{A}^{+}\right) \leqslant 1 \end{cases}$$

for given nonnegative definite matrices $\mathbf{A}$, $\mathbf{B}$, and the inherent relationships

$$\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X} \leqslant \mathbf{X}'\mathbf{W}^{+}\mathbf{X}, \quad P_{\mathbf{F}'}\mathbf{W} = \mathbf{W}P_{\mathbf{F}'} \quad \Leftrightarrow \quad \mathscr{R}\left(\mathbf{W}\mathbf{F}'\right) \subseteq \mathscr{R}\left(\mathbf{F}'\right)$$

one can justify that

Corollary 2.2. *Assume* θ *is nonnegative quadratic estimable under model* $\mathscr{L}_T$. *Then the statements below hold*:

(a) *If* $\mathscr{R}\left(\mathbf{X},\ \mathbf{W}\mathbf{F}'\right) \subseteq \mathscr{R}\left(\mathbf{F}'\right)$, *then* $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ *if and only if* $\hat{h}\left(s - \tilde{s}\right) = 0$,
(b) *If* $\mathscr{R}\left(\mathbf{F}'\right) = \mathscr{R}\left(\mathbf{W}\right)$, *then* $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$.

The wonderful result (b) derived in Corollary 2.2 reveals a fact, from something more than intuitional impression, that if $\mathscr{R}\left(\mathbf{F}'\right) = \mathscr{R}\left(\mathbf{W}\right)$, then the transformed model contains all information of the original model in a sense.

2.2. Nonnegative quadratic-biased estimation and quadratic sufficiency

As argued above that $\hat{\theta}$ and $\tilde{\theta}$ are unacceptable in practice when θ is not an nonnegative quadratic estimable function, or, when $\hat{h}$ and $\tilde{h}$ are negative, under this setting we consider the notion of quadratic sufficiency in the case $\mathscr{C} = \mathscr{C}_2$, the class of nonnegative quadratic-biased estimation (see below) for θ. To demonstrate our conclusions, we first state some preliminary knowledge concisely in the following. Put

$$\hat{\theta}\left(\widehat{\mathbf{C}}, \hat{c}\right) = \hat{\beta}'\widehat{\mathbf{C}}\hat{\beta} + \hat{c}\left(y - \mathbf{X}\hat{\beta}\right)'\mathbf{W}^{+}\left(y - \mathbf{X}\hat{\beta}\right),$$
$$\tilde{\theta}\left(\widetilde{\mathbf{C}}, \tilde{c}\right) = \tilde{\beta}'\widetilde{\mathbf{C}}\tilde{\beta} + \tilde{c}\left(y - \mathbf{X}\tilde{\beta}\right)'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\left(y - \mathbf{X}\tilde{\beta}\right)$$

with $\widehat{\mathbf{C}}, \widetilde{\mathbf{C}} \in \mathbb{R}_p^{\geqslant}$ and $\hat{c}, \tilde{c} \geqslant 0$ (we can concentrate our mind on $\mathbb{R}_p^{\geqslant}$ since it is the convex cone of $p \times p$ symmetric matrices). The following Definition 2.2 concerns NN-BQB estimator, with above form, of θ by Lemma 2.1; e.g., cf. [6]. For another method to find the minimum-biased estimator in the class of all quadratic forms $y'\mathbf{A}y$, one can see Gnot et al. [7].

Definition 2.2. We say $\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}, \hat{c}_{\mathbf{H}}\right)$ is an NN-BQB estimator for θ under model $\mathscr{L}$ if the pair $\left(\widehat{\mathbf{C}}_{\mathbf{H}}, \hat{c}_{\mathbf{H}}\right)$ solves the following problem:

$$\min_{\widehat{\mathbf{C}} \in \mathbb{R}_{p,}^{\geqslant}\, \hat{c} \geqslant 0} \operatorname{tr}\left\{\left(\mathbf{H} - \widehat{\mathbf{C}}\right)^2\right\} + \left[\hat{c} - h + \operatorname{tr}\left\{\widehat{\mathbf{C}}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+} - \widehat{\mathbf{C}}P_{\mathbf{X}'}\mathbf{U}P_{\mathbf{X}'}\right\}\right]^2. \tag{2.4}$$

If $\hat{h} < 0$, then the problem in search of the solution to (2.4) reduces to minimizing $f_{\mathbf{H}}\left(\widehat{\mathbf{C}}\right)$ with respect to $\widehat{\mathbf{C}}$ varying over $\mathbb{R}_p^{\geqslant}$ since we can restrict our considerations to $\hat{\beta}'\widehat{\mathbf{C}}\hat{\beta}$ (cf. [6]), where

$$f_{\mathbf{H}}\left(\widehat{\mathbf{C}}\right) = \operatorname{tr}\left\{\left(\mathbf{H} - \widehat{\mathbf{C}}\right)^2\right\} + \left(h - \operatorname{tr}\left\{\widehat{\mathbf{C}}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ - \widehat{\mathbf{C}}P_{\mathbf{X}'}\mathbf{U}P_{\mathbf{X}'}\right\}\right)^2.$$

We now adopt the Lagrange multiplier method. Let $\mathrm{L}_{\mathbf{H}}\left(\widehat{\mathbf{C}}, \mathbf{\Lambda}\right) = f_{\mathbf{H}}\left(\widehat{\mathbf{C}}\right) - 2\operatorname{tr}\left(\widehat{\mathbf{C}}\mathbf{\Lambda}\right)$ be the Lagrange function, where $\mathbf{\Lambda}$ is some nonnegative Lagrange multiplier matrix. Then there exists a nonnegative definite matrix $\mathbf{\Lambda}_{\mathbf{H}}$ such that for any solution $\widehat{\mathbf{C}}_{\mathbf{H}}$ to minimizing $f_{\mathbf{H}}\left(\widehat{\mathbf{C}}\right)$, $f_{\mathbf{H}}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right) = \min_{\widehat{\mathbf{C}}\in\mathbb{R}_p^{\geqslant}} L_{\mathbf{H}}\left(\widehat{\mathbf{C}}, \Lambda_{\mathbf{H}}\right)$ and

$$\operatorname{tr}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\Lambda_{\mathbf{H}}\right) = 0. \tag{2.5}$$

By means of standard formulas for partially derivatives of trace functions and the fact that $\mathbf{A} = \mathbf{B} \Leftrightarrow 2\mathbf{A} - \operatorname{diag}(\mathbf{A}) = 2\mathbf{B} - \operatorname{diag}(\mathbf{B})$ for given $\mathbf{A}, \mathbf{B} \in \mathbb{R}_n^{s}$, where $\operatorname{diag}(\mathbf{A})$ denotes a diagonal matrix obtained from $\mathbf{A}$ by replacing all the off-diagonal elements of $\mathbf{A}$ by zeros, we obtain that the gradient of $L_{\mathbf{H}}\left(\widehat{\mathbf{C}}, \mathbf{\Lambda}_{\mathbf{H}}\right)$, $L_{\mathbf{H}}\left(\widehat{\mathbf{C}}, \mathbf{\Lambda}_{\mathbf{H}}\right) \Big/ \partial\widehat{\mathbf{C}}$, with respect to $\widehat{\mathbf{C}}$ vanishes at $\widehat{\mathbf{C}}_{\mathbf{H}}$ if and only if the following relationship holds:

$$\begin{aligned}\mathbf{\Lambda}_{\mathbf{H}} = {} & \widehat{\mathbf{C}}_{\mathbf{H}} - \mathbf{H} + \left(\operatorname{tr}\left\{\widehat{\mathbf{C}}_{\mathbf{H}}\left[\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ - P_{\mathbf{X}'}\mathbf{U}P_{\mathbf{X}'}\right]\right\} - h\right) \\ & \times\left[\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+ - P_{\mathbf{X}'}\mathbf{U}P_{\mathbf{X}'}\right].\end{aligned} \tag{2.6}$$

If $\tilde{h}$ is negative, then NN-BQB estimation for θ under model $\mathscr{L}_T$ could be discussed with similar fashion and be omitted here.

Generally speaking, it is not easy to find an explicit solution to (2.5) and (2.6). For the case $\mathbf{V} = \mathbf{I}_n$ (and therefore taking $\mathbf{U} = \mathbf{0}$ is a suitable and simple choice), however, it has been done if $\mathbf{H}$ and $\mathbf{X}'\mathbf{X}$ commute, which is fulfilled in many practical situations, and a procedure leading to the solution to (2.5) and (2.6) without a commutativity assumption has been described in Gnot et al. [6].

Let us now pay attention to $\mathscr{S}_{(\mathscr{C}_2)}$ since we mainly focus our mind on the problem of quadratic sufficiency. In the following, we assume $\widehat{\mathbf{C}}_{\mathbf{H}}$ is a solution to (2.5) and (2.6) and thereby $\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right) = \hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}, 0\right) = y'\mathbf{A}_{\mathbf{H}}y$ is an NN-BQB estimator of θ under model $\mathscr{L}$ in the sense of Definition 2.2; Similarly, we assume $\tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right) = \tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}, 0\right) = y'\mathbf{B}_{\mathbf{H}}y$ is an NN-BQB estimator of θ under model $\mathscr{L}_T$, with notations

$$\begin{aligned}\mathbf{A}_{\mathbf{H}} &= \mathbf{W}^+\mathbf{X}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+\widehat{\mathbf{C}}_{\mathbf{H}}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+\mathbf{X}'\mathbf{W}^+, \\ \mathbf{B}_{\mathbf{H}} &= \mathbf{F}'\widetilde{\mathbf{W}}^+\mathbf{F}\mathbf{X}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^+\mathbf{F}\mathbf{X}\right)^+\widetilde{\mathbf{C}}_{\mathbf{H}}\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^+\mathbf{F}\mathbf{X}\right)^+\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^+\mathbf{F}.\end{aligned}$$

According to (2.2) one can verify that

$$\begin{aligned}\operatorname{Cov}\left[\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right), \tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)\right] = {} & 2\sigma^4\operatorname{tr}\left\{P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+\widehat{\mathbf{C}}_{\mathbf{H}}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+\right\} - 4\sigma^4 \\ & \times\operatorname{tr}\left\{P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'}\mathbf{U}P_{\mathbf{X}'}\widehat{\mathbf{C}}_{\mathbf{H}}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+\right\} + 2\sigma^4\operatorname{tr}\left\{P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'}\mathbf{U}P_{\mathbf{X}'}\widehat{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'}\mathbf{U}\right\} \\ & +4\sigma^2\beta'P_{\mathbf{X}'}\widehat{\mathbf{C}}_{\mathbf{H}}\left(\mathbf{X}'\mathbf{W}^+\mathbf{X}\right)^+P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'}\beta - 4\sigma^2\beta'P_{\mathbf{X}'}\widehat{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'}\mathbf{U}P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'}\beta.\end{aligned}$$

Observe that replacing $\left(\mathbf{F},\ \widetilde{\mathbf{C}}_{\mathbf{H}}\right)$ with corresponding $\left(\mathbf{I}_n,\ \widehat{\mathbf{C}}_{\mathbf{H}}\right)$ in the expression of Cov $\left[\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right),\ \tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)\right]$ gives $\mathscr{D}\left[\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right)\right]$, and replacing $\left(\mathbf{X},\ \widehat{\mathbf{C}}_{\mathbf{H}}\right)$ by $\left(\mathbf{FX},\ \widetilde{\mathbf{C}}_{\mathbf{H}}\right)$ in the presentation of $\mathscr{D}\left[\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right)\right]$ further yields $\mathscr{D}\left[\tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)\right]$.

The following theorem gives some results concerning $\mathscr{S}_{(\mathscr{C}_2)}$.

Theorem 2.3. $\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right)$ *and* $\tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)$ *are identical if and only if the following two conditions hold simultaneously*:

- $P_{\mathbf{X}'}\widehat{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'} = P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'}$,
- $\left[\left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{FX}\right)^{+} - P_{\mathbf{X}'\mathbf{F}'}\left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}P_{\mathbf{X}'\mathbf{F}'}\right]\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'} = \mathbf{0}$.

When the above two conditions are satisfied, $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_2)}$. *Further*, *if NN-BQB estimate is essentially unique*, *then the two conditions are sufficient* (*truly*) *and yet necessary*.

Proof. For convenience, we write

$$\begin{aligned}
&\mathbf{A} = P_{\mathbf{X}'}\widehat{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'}, \quad \widetilde{\mathbf{A}} = P_{\mathbf{X}'\mathbf{F}'}\widetilde{\mathbf{C}}_{\mathbf{H}}P_{\mathbf{X}'\mathbf{F}'},\\
&\mathbf{B} = \left(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}\right)^{+}, \quad \widetilde{\mathbf{B}} = \left(\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{FX}\right)^{+},\\
&\mathbf{C} = P_{\mathbf{X}'}\mathbf{U}P_{\mathbf{X}'}, \quad \widetilde{\mathbf{C}} = P_{\mathbf{X}'\mathbf{F}'}\mathbf{U}P_{\mathbf{X}'\mathbf{F}'}.
\end{aligned}$$

It follows that $\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right)$ and $\tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)$ are identical iff $\mathscr{E}\left[\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right) - \tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)\right]=0$ and $\mathscr{D}\left[\hat{\theta}\left(\widehat{\mathbf{C}}_{\mathbf{H}}\right) - \tilde{\theta}\left(\widetilde{\mathbf{C}}_{\mathbf{H}}\right)\right]=0$, or equivalently, the following hold simultaneously:

(1) $\operatorname{tr}\left[\mathbf{A}\left(\mathbf{B}-\mathbf{C}\right)\right] = \operatorname{tr}\left[\widetilde{\mathbf{A}}\left(\widetilde{\mathbf{B}}-\widetilde{\mathbf{C}}\right)\right]$,
(2) $\mathbf{A} = \widetilde{\mathbf{A}}$,
(3) $\operatorname{tr}(\mathbf{ABAB}) + \operatorname{tr}(\mathbf{ACAC}) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}\right) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\mathbf{C}\widetilde{\mathbf{A}}\mathbf{C}\right) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\mathbf{CAB}\right) + \operatorname{tr}\left(\mathbf{AB}\widetilde{\mathbf{A}}\mathbf{C}\right)$ $+\operatorname{tr}\left(\widetilde{\mathbf{A}}\mathbf{BAC}\right) + \operatorname{tr}\left(\mathbf{AC}\widetilde{\mathbf{A}}\mathbf{B}\right) = \operatorname{tr}(\mathbf{ACAB}) + \operatorname{tr}(\mathbf{ABAC}) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\mathbf{C}\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}\right) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}\widetilde{\mathbf{A}}\mathbf{C}\right) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\mathbf{BAB}\right)$ $+ \operatorname{tr}\left(\mathbf{AB}\widetilde{\mathbf{A}}\mathbf{B}\right) + \operatorname{tr}\left(\widetilde{\mathbf{A}}\mathbf{CAC}\right) + \operatorname{tr}\left(\mathbf{AC}\widetilde{\mathbf{A}}\mathbf{C}\right)$,
(4) $\mathbf{ABA} + \widetilde{\mathbf{A}}\widetilde{\mathbf{B}}\widetilde{\mathbf{A}} + \mathbf{AC}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{CA} = \mathbf{ACA} + \widetilde{\mathbf{A}}\mathbf{C}\widetilde{\mathbf{A}} + \mathbf{AB}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{BA}$.

By the above four representations we can readily justify the conclusion and thus complete the proving. □

3. Further results and simulation

In this section, we pose a solvable practical problem, extend partially the notion of quadratic sufficiency to multivariate case, and conduct a simulated example.

3.1. A practical problem

As we can see, a problem which needs solving in practice is that if the observation vector y cannot be available while $\mathbf{F}_1 y$ and $\mathbf{F}_2 y$ are two obtainable vectors, which one we would choose. Write

$$\tilde{\tilde{\theta}}_i = \tilde{\beta}'\mathbf{H}\tilde{\beta} + \tilde{\sigma}^2\left[h - \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{F}_i'\left(\mathbf{F}_i\mathbf{W}\mathbf{F}_i'\right)^{+}\mathbf{F}_i\mathbf{X}\right)^{+} - \mathbf{HU}\right\}\right], \quad i = 1, 2.$$

We state our conclusion, which can be proved similarly to Theorem 2.1, in the following:

Theorem 3.1. *Assume θ is nonnegative quadratic estimable under models*

$$\left\{\mathbf{F}_i y,\ \mathbf{F}_i\mathbf{X}\beta,\ \sigma^2\mathbf{F}_i\mathbf{V}\mathbf{F}_i'\right\}, \quad i = 1, 2$$

and

$$h > \operatorname{tr}\left\{\mathbf{H}\left(\mathbf{X}'\mathbf{F}_i'\left(\mathbf{F}_i\mathbf{W}\mathbf{F}_i'\right)^+\mathbf{F}_i\mathbf{X}\right)^+ - \mathbf{H}\mathbf{U}\right\}.$$

Then $\tilde{\tilde{\theta}}_1$ is superior over $\tilde{\tilde{\theta}}_2$ in the sense that $\mathscr{D}\left(\tilde{\tilde{\theta}}_1\right) \leqslant \mathscr{D}\left(\tilde{\tilde{\theta}}_2\right)$ if and only if

- $\mathbf{H}\left(\mathbf{X}'\mathbf{F}_1'\left(\mathbf{F}_1\mathbf{W}\mathbf{F}_1'\right)^+\mathbf{F}_1\mathbf{X}\right)^+\mathbf{H} \leqslant \mathbf{H}\left(\mathbf{X}'\mathbf{F}_2'\left(\mathbf{F}_2\mathbf{W}\mathbf{F}_2'\right)^+\mathbf{F}_2\mathbf{X}\right)^+\mathbf{H}$,
- $\operatorname{rk}(\mathbf{F}_1) - \operatorname{rk}(\mathbf{F}_1\mathbf{X}) \geqslant \operatorname{rk}(\mathbf{F}_2) - \operatorname{rk}(\mathbf{F}_2\mathbf{X})$.

Similarly, we can investigate other cases and we omit them here.

3.2. Multivariate case

As the supplement and an application of our main results and an extension of Baksalary and Drygas [1], we now consider the multivariate case as follows. Let $\mathbf{A} \otimes \mathbf{B}$ stand for the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. For a multivariate linear model, denoted by

$$\mathscr{L}^{(M)} = \left\{\mathbf{Y},\ \mathbf{X}\mathscr{B},\ \mathbf{I}_q \otimes \sigma^2\mathbf{V}\right\}, \tag{3.1}$$

in which $\mathbf{Y} = \left(y_1, \ldots, y_q\right)$ refers to an $n \times q$ normally distributed random matrix of observations, $y_1, \ldots, y_q$ are independent and with identical covariance matrix $\sigma^2\mathbf{V}$, $\mathbf{X} \in \mathbb{R}_{n,p}$ and $\mathbf{V} \in \mathbb{R}_n^{\geqslant}$ are supposed to be known, while $\mathscr{B} \in \mathbb{R}_{p,q}$ and $\sigma^2 > 0$ are unknown parameters. Being similar to Section 1, we consider the transformed model, given by

$$\mathscr{L}_T^{(M)} = \left\{\mathbf{F}\mathbf{Y},\ \mathbf{F}\mathbf{X}\mathscr{B},\ \mathbf{I}_q \otimes \sigma^2\mathbf{F}\mathbf{V}\mathbf{F}'\right\} \tag{3.2}$$

with $\mathbf{F} \in \mathbb{R}_{m,n}$, $m \geqslant p$. Now we generalize the concept of quadratic sufficiency defined by Baksalary and Drygas [1] in the following. Let

$$h^* = \operatorname{rk}\left(\mathbf{I}_q \otimes \mathbf{V}, \mathbf{I}_q \otimes \mathbf{X}\right) - \operatorname{rk}\left(\mathbf{I}_q \otimes \mathbf{X}\right).$$

Definition 3.1. For the models $\mathscr{L}^{(M)}$ and $\mathscr{L}_T^{(M)}$, the transformation $\mathbf{FY}$ is said quadratically sufficient if there is a symmetric matrix $\mathbf{G}$ such that $\operatorname{tr}\left(\mathbf{Y}'\mathbf{F}'\mathbf{G}\mathbf{F}\mathbf{Y}\right)$ is the best quadratic unbiased estimator for $h^*\sigma^2$. We write $\mathbf{FY} \in \mathscr{S}^{(M)}$ if $\mathbf{FY}$ is quadratically sufficient.

The following theorem gives some results concerning $\mathscr{S}^{(M)}$.

Theorem 3.2. $\mathbf{FY} \in \mathscr{S}^{(M)}$ *iff* $\operatorname{rk}(\mathbf{V},\ \mathbf{X}) - \operatorname{rk}(\mathbf{X}) = \operatorname{rk}\left(\mathbf{F}\mathbf{V}\mathbf{F}',\ \mathbf{F}\mathbf{X}\right) - \operatorname{rk}(\mathbf{F}\mathbf{X})$.

Proof. Write models $\mathscr{L}^{(M)}$ and $\mathscr{L}_T^{(M)}$ as

$$\left\{y, \left(\mathbf{I}_q \otimes \mathbf{X}\right)\beta,\ \mathbf{I}_q \otimes \sigma^2\mathbf{V}\right\}, \quad \left\{\mathbf{F}y, \left(\mathbf{I}_q \otimes \mathbf{F}\right)\left(\mathbf{I}_q \otimes \mathbf{X}\right)\beta,\ \mathbf{I}_q \otimes \sigma^2\mathbf{F}\mathbf{V}\mathbf{F}'\right\},$$

respectively, in which $y = \mathrm{Vec}\,(\mathbf{Y})$ and $\beta = \mathrm{Vec}\,(\mathscr{B})$, where the vector $\mathrm{Vec}\,(\mathbf{A})$ is obtained from the matrix $\mathbf{A}$ by stacking its columns one underneath the other. Notice that

$$\mathrm{tr}\left(\mathbf{Y}'\mathbf{F}'\mathbf{G}\mathbf{F}\mathbf{Y}\right) = y'\left(\mathbf{I}_q \otimes \mathbf{F}'\mathbf{G}\mathbf{F}\right) y = y'\left(\mathbf{I}_q \otimes \mathbf{F}\right)'\left(\mathbf{I}_q \otimes \mathbf{G}\right)\left(\mathbf{I}_q \otimes \mathbf{F}\right) y \tag{3.3}$$

in terms of the standard formula (e.g., cf. [18]) for trace of matrix that

$$\mathrm{tr}\,(\mathbf{ABC}) = \left(\mathrm{Vec}\left(\mathbf{A}'\right)\right)'(\mathbf{I} \otimes \mathbf{B})\ \mathrm{Vec}\,(\mathbf{C})\,.$$

Eq. (3.3) would imply that $\mathbf{FY} \in \mathscr{S}^{(\mathbf{M})}$ if and only if $\left(\mathbf{I}_q \otimes \mathbf{F}\right) y$ being quadratically sufficient in the sense of Baksalary and Drygas [1]. By virtue of Corollary 2.1 and the facts that $\mathrm{rk}\,(\mathbf{A} \otimes \mathbf{B}) = \mathrm{rk}\,(\mathbf{A})\,\mathrm{rk}\,(\mathbf{B})$ and

$$\begin{aligned}\mathrm{rk}\,(\mathbf{I} \otimes \mathbf{X},\ \mathbf{I} \otimes \mathbf{V}) &= \mathrm{rk}\begin{pmatrix} \mathbf{X} & & & \mathbf{V} & & \\ & \ddots & & & \ddots & \\ & & \mathbf{X} & & & \mathbf{V} \end{pmatrix} \\ &= \mathrm{rk}\begin{pmatrix} (\mathbf{X},\ \mathbf{V}) & & \\ & \ddots & \\ & & (\mathbf{X},\ \mathbf{V}) \end{pmatrix} = \mathrm{rk}\,[\mathbf{I} \otimes (\mathbf{X},\ \mathbf{V})]\end{aligned}$$

one could readily justify the conclusion. $\square$

By Theorem 3.2, it should be noticed that the necessary and sufficient condition for $\mathbf{FY} \in \mathscr{S}^{(M)}$ is independent of the scalar q.

3.3. A simulation study

In this subsection, a simulation study was conducted. For models $\mathscr{L}$ and $\mathscr{L}_T$ with

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 3 \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}',$$

we take $h = 0.7$, $\sigma^2 = 0.09$ and

$$\mathbf{U} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 0.2500 & 0.2000 \\ 0.2000 & 0.2500 \end{pmatrix}, \quad \beta = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

It is clear that $\mathscr{R}\,(\mathbf{X}) \subseteq \mathscr{R}\left(\mathbf{V} + \mathbf{XUX}'\right)$ and $\mathscr{R}\,(\mathbf{H}) \subseteq \mathscr{R}\left(\mathbf{X}'\right)$. In addition, the value of $\theta = \beta'\mathbf{H}\beta + h\sigma^2$ is equal to 11.1130, and

$$\widehat{\mathbf{A}} + \hat{b}\widehat{\mathbf{B}} = \begin{pmatrix} 0.0521 & 0.0479 & 0.0021 & 0.0062 \\ 0.0479 & 0.1021 & 0.0479 & -0.0062 \\ 0.0021 & 0.0479 & 0.0521 & 0.0062 \\ 0.0062 & -0.0062 & 0.0062 & 0.0187 \end{pmatrix},$$

$$\widetilde{\mathbf{A}} + \tilde{b}\widetilde{\mathbf{B}} = \begin{pmatrix} 0.1333 & 0.0417 & 0.0833 & -0.0417 \\ 0.0417 & 0.0333 & 0.0417 & -0.0333 \\ 0.0833 & 0.0417 & 0.1333 & -0.0417 \\ -0.0417 & -0.0333 & -0.0417 & 0.0333 \end{pmatrix}.$$

The following Table 1 lists the estimates for θ obtained via $\hat{\theta}$ and $\tilde{\theta}$.

Table 1
Estimates for θ with random outputs of y

$\hat{\theta}$	$\tilde{\theta}$	$\hat{\theta}$	$\tilde{\theta}$	$\hat{\theta}$	$\tilde{\theta}$	$\hat{\theta}$	$\tilde{\theta}$
11.7855	11.8520	11.1203	10.7488	9.56060	9.87180	11.4357	11.6741
11.7448	11.6975	11.3278	11.2417	10.7829	10.5625	11.3616	11.7646
12.3717	12.1100	10.2087	10.2318	9.24410	9.16610	10.9183	10.9882
10.4737	10.5446	10.5971	10.5623	12.2174	11.9867	10.5472	10.5575
11.4109	11.5208	13.6636	13.4735	9.70510	9.48140	11.6134	10.6075

4. Applications

We consider a linear regression model denoted by $\mathscr{L}$ with the covariance matrix $\mathbf{V}$ being compound symmetric, that is,

$$\mathbf{V} = (1-\rho)\,\mathbf{I}_n + \rho \mathbf{1}_n \mathbf{1}_n' = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

for given $\rho \in [0, 1)$. Notice that the condition $0 \leqslant \rho < 1$ ensures that $\mathbf{V}$ is positive definite and thus taking $\mathbf{U} = \mathbf{0}$ is a suitable and simple choice. Let us now apply our main results to the problem in estimating $\mathscr{E}\left(y'y\right)$, which has the form of θ with $\mathbf{H} = \mathbf{X}'\mathbf{X}$ and $h = n$. We assume $\hat{h} = n - \mathrm{tr}[\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'] > 0$ in the following.

Theorem 4.1. *For the models $\mathscr{L}$ and $\mathscr{L}_T$ with compound symmetric covariance matrix*, $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ *if and only if* $\mathscr{N}(\mathbf{F}) \subseteq \mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}\left(\mathbf{V}\mathbf{F}'\right)$, *where the symbol* $\mathscr{N}(\mathbf{F})$ *refers to the null space of* $\mathbf{F}$, *i.e.*, $\mathscr{N}(\mathbf{F}) = \mathscr{R}\left(\mathbf{F}'^{\perp}\right) = \mathscr{R}\left(\mathbf{F}'\right)^{\perp}$.

Proof. From Theorem 2.2 and Corollary 2.1, $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ if and only if

$$\left(\mathbf{X}'\mathbf{F}'\left(\mathbf{F}\mathbf{V}\mathbf{F}'\right)^{-}\mathbf{F}\mathbf{X}\right)\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{X} = \mathbf{X}'\mathbf{X},$$
$$\mathscr{R}\left(\mathbf{V}\mathbf{X}^{\perp}\right) = \mathscr{R}\left(\mathbf{V}\mathbf{F}'\left(\mathbf{F}\mathbf{X}\right)^{\perp}\right),$$

or equivalently, $\mathbf{X}'\mathbf{F}'\left(\mathbf{F}\mathbf{V}\mathbf{F}'\right)^{-}\mathbf{F}\mathbf{X} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ and $\mathscr{R}\left(\mathbf{X}^{\perp}\right) = \mathscr{R}\left(\mathbf{F}'\left(\mathbf{F}\mathbf{X}\right)^{\perp}\right)$. Observe that the first condition is equivalent to

$$\left(\mathbf{I}_n - P_{\mathbf{V}^{\frac{1}{2}}\mathbf{F}'}\right)\mathbf{V}^{-\frac{1}{2}}\mathbf{X} = \mathbf{0}$$

in terms of the relationship $\mathbf{X}'\mathbf{F}'\widetilde{\mathbf{W}}^{+}\mathbf{F}\mathbf{X} \leqslant \mathbf{X}'\mathbf{W}^{+}\mathbf{X}$, or equivalently, $\mathscr{R}\left(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\right) \subseteq \mathscr{R}\left(\mathbf{V}^{\frac{1}{2}}\mathbf{F}'\right)$, which is further equivalent to $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}\left(\mathbf{V}\mathbf{F}'\right)$; the second is satisfied if and only if $\mathscr{R}\left(\mathbf{X}^{\perp}\right) \subseteq \mathscr{R}\left(\mathbf{F}'\right)$, or equivalently, $\mathscr{N}(\mathbf{F}) \subseteq \mathscr{R}(\mathbf{X})$. The proving is thus completed. □

It can be seen that if $\mathbf{F}$ is full column rank, then $\mathscr{L} \Rightarrow \mathscr{L}_T$ by pre-multiplying $\mathscr{L}$ by $\mathbf{F}$ and $\mathscr{L}_T \Rightarrow \mathscr{L}$ by pre-multiplying $\mathscr{L}_T$ by $\mathbf{F}^{+}$ since $\mathbf{F}^{+}\mathbf{F} = \mathbf{I}_n$, that is to say, $\mathscr{L}_T$ contains all information of $\mathscr{L}$ and vice versa. The following corollaries indicate $\mathbf{F}$ being full column rank is sufficient (truly in the above sense) and yet necessary for models $\mathscr{L}$ and $\mathscr{L}_T$ to coincide with each other for some special cases. On the other hand, the problem of the notion of sufficiency (such as linear sufficiency, linear error-sufficiency, quadratic sufficiency) is mentioned usually when y is not available (but $\mathbf{F}y$ is obtainable). However, it also may be used to simplify calculating by choosing suitable $\mathbf{F}$ when $\hat{\theta}$ is complicated and hard to calculate. This idea is not only mathematically interesting but also noticeable in practice. Corollaries 4.1 and 4.2 give a partial response.

Corollary 4.1. *For the models $\mathscr{L}$ and $\mathscr{L}_T$ with compound symmetric covariance matrix, if $\mathbf{F}'$ has one column as the unit vector. Then $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ if and only if* $\mathrm{rk}\,(\mathbf{F}) = n$.

Proof. Without loss of generality, $\mathbf{F}' = (\mathbf{1}_n, \mathbf{F}_2)$. It follows that $P_{\mathbf{F}'}\mathbf{V} = \mathbf{V}P_{\mathbf{F}'}$ (e.g., cf. [17],), or equivalently, $\mathscr{R}\left(\mathbf{V}\mathbf{F}'\right) \subseteq \mathscr{R}\left(\mathbf{F}'\right)$. This fact combining Theorem 4.1 would imply

$$\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)} \quad \Leftrightarrow \quad \mathscr{R}\left(\mathbf{X},\, \mathbf{X}^{\perp}\right) \subseteq \mathscr{R}\left(\mathbf{F}'\right) \quad \Leftrightarrow \quad \mathrm{rk}\,(\mathbf{F}) = n,$$

which completes the proving. $\square$

Similarly, we have

Corollary 4.2. *For the models $\mathscr{L}$ and $\mathscr{L}_T$ with compound symmetric covariance matrix, assume $\mathbf{X}$ has one column as the unit vector. Then $\mathbf{F}y \in \mathscr{S}_{(\mathscr{C}_1)}$ if and only if* $\mathrm{rk}\,(\mathbf{F}) = n$.

5. Concluding remarks

In this article the problem of nonnegative estimation of $\beta'\mathbf{H}\beta + h\sigma^2$ is discussed under models $\mathscr{L}$ and $\mathscr{L}_T$. A generalization of quadratic sufficiency is considered. Actually, our main results are partially supported by Eq. (2.2), which holds under assumption yet in the quasi-normal case; e.g., cf. [12], see also ([18], Lemma 1.1, p. 144).

Thus, we can establish our conclusions under quasi-normality assumption with similar fashion by rewriting model $\mathscr{L}$ as

$$y = \mathbf{X}\beta + \mathbf{C}\varepsilon,$$

where $\mathbf{C}\mathbf{C}' = \mathbf{V}$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_v)'$ (here v refers to the rank of $\mathbf{V}$) with $\mathscr{E}\,(\varepsilon_i) = 0$, $\mathscr{E}\left(\varepsilon_i^2\right) = \sigma^2$, $\mathscr{E}\left(\varepsilon_i^3\right) = 0$, $\mathscr{E}\left(\varepsilon_i^4\right) = 3\sigma^4$, i.e., the components of ε behave up to their moments up to order four as independent normally distributed random variables.

Acknowledgments

The authors are grateful to the Editor and the four anonymous referees for their very careful reading and constructive comments which would improve the paper substantially. Thanks also to Prof. Dao-Yuan Zhu (Southeast University, China) for his beneficial suggestions, and to the Journal Manager, Prof. Oscar Sosa, for time and patience with helpful aspects in the whole process of dealing with this paper.

Appendix A. Essentially unique BQU estimate

Here, we show that if θ is nonnegative quadratic estimable in the model $\mathscr{L}$, then $\hat{\theta}$ is the essentially unique BQU estimator of θ under normality assumption with respect to MSE criterion. Actually, using techniques of the inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{AB})$ in $\mathbb{R}_n^{\geqslant}$ is helpful for us to establish the same conclusion under quasi-normality assumption (see Section 5). We now offer an alternative proof as follows. It should be mentioned that the approach is a version of the well-known theorem of Lehmann–Scheffé, which says that an unbiased estimator is optimal if it is uncorrelated with any unbiased estimator of zero, in some sense since the set of estimators here is restricted to the quadratic estimators. Actually, denote

$$\widehat{\mathbf{D}} = \widehat{\mathbf{A}} + \hat{b}\widehat{\mathbf{B}}$$

and let $y'\mathbf{D}y$ be an alternative quadratic unbiased estimator for θ, or equivalently, $\mathbf{X}'\mathbf{DX} = \mathbf{H}$ and $\operatorname{tr}(\mathbf{DV}) = h$. First of all, writing $\mathbf{C} = \mathbf{D} - \widehat{\mathbf{D}}$ yields $\mathbf{X}'\mathbf{CX} = 0$ and $\operatorname{tr}(\mathbf{CV}) = 0$, and therefore $\operatorname{tr}(\mathbf{CW}) = 0$. Further, $\operatorname{tr}\left(\widehat{\mathbf{D}}\mathbf{VCV}\right) = 0$ and $\mathbf{X}'\widehat{\mathbf{D}}\mathbf{VCX} = \mathbf{0}$ by direct operations. So we obtain

$$\operatorname{Cov}\left(y'\widehat{\mathbf{D}}y, y'\mathbf{C}y\right) = 2\sigma^4 \operatorname{tr}\left(\widehat{\mathbf{D}}\mathbf{VCV}\right) + 4\sigma^2\beta'\mathbf{X}'\widehat{\mathbf{D}}\mathbf{VCX}\beta = 0.$$

This fact would imply that

$$\mathscr{D}\left(y'\mathbf{D}y\right) = \mathscr{D}\left(y'\widehat{\mathbf{D}}y\right) + \mathscr{D}\left(y'\mathbf{C}y\right) \geqslant \mathscr{D}\left(y'\widehat{\mathbf{D}}y\right),$$

which combining the unbiasedness of $\hat{\theta}$ for θ completes the proving of optimality. Clearly, the equality occurs in the above expression if and only if $\mathscr{D}\left(y'\mathbf{C}y\right) = 0$, or equivalently,

$$\operatorname{tr}(\mathbf{CVCV}) = 0, \quad \mathbf{X}'\mathbf{CVCX} = \mathbf{0}. \tag{A.1}$$

Together $\mathbf{X}'\mathbf{CX} = \mathbf{0}$, $\operatorname{tr}(\mathbf{CV}) = 0$ and Eq. (A.1), we have $\mathbf{WCW} = \mathbf{0}$, and vice versa. Note that $\mathbf{WCW} = \mathbf{0}$ means $y'\mathbf{C}y = 0$ almost surely and thereby $y'\mathbf{D}y = y'\widehat{\mathbf{D}}y + y'\mathbf{C}y = y'\widehat{\mathbf{D}}y$ with probability one. The uniqueness is also established.

References

[1] J.K. Baksalary, H.G. Drygas, A note on the concepts of sufficiency in the general Gauss–Markov model—a coodinate-free aproach, Mathematische Schriften Kassel, Vordruck-Reihe des Fachbereichs 17 der Gesamthochschule Kassel, Preprint No. 92/2, 1992.

[2] J.K. Baksalary, T. Mathew, Linear sufficiency and completeness in an incorrectly specified general Gauss–Markov model, Sankhya A 48 (1986) 169–180.

[3] H. Drygas, Sufficiency and completeness in general Gauss–Markoff model, Sankhya A 45 (1983) 88–98.

[4] H. Drygas, Linear sufficiency and some applications in multilinear estimation, J. Multivariate Anal. 16 (1985) 71–84.

[5] H. Drygas, Linear and quadratic sufficiency—the linear Rao–Blackwell Theorem, in: S.K. Basu, B.K. Sinha (Eds.), Proceedings of the Calcutta Statistical Meeting, 27-12-1991–1-1-1992, Norosa Publishing House, New Delhi, 1993, pp. 121–125.

[6] S. Gnot, G. Trenkler, R. Zmyślony, Nonnegative minimum biased quadratic estimation in the linear regression models, J. Multivariate Anal. 54 (1995) 113–125.

[7] S. Gnot, M. Grzadziel, Nonnegative minimum biased quadratic estimation in mixed linear models, J. Multivariate Anal. 80 (2002) 217–233.

[9] J. Groß, A note on the concepts of linear and quadratic sufficiency, J. Statist. Plann. Inference 70 (1998) 69–76.

[11] B. Heiligers, A. Markiewicz, Linear sufficiency and admissibility in restricted linear models, Statist. Probab. Lett. 30 (1996) 105–111.

[12] P.L. Hsu, On the best unbiased quadratic estimation of heteroscedastic variances, Math. Oper. Statist. Statist. 1 (1938) 27–44.
[13] W.C. Ip, H. Wong, J.S. Liu, Sufficient and admissible estimators in general multivariate linear model, J. Statist. Plann. Inference 135 (2005) 371–383.
[15] A. Markiewicz, Comparison of linear restricted models with respect to the validity of admissible and linearly sufficient estimators, Statist. Probab. Lett. 38 (1998) 347–354.
[16] J. Müeller, Sufficiency and completeness in the linear model, J. Multivariate Anal. 21 (1987) 312–323.
[17] C.R. Rao, H. Toutenburg, Linear Models: Least squares and Alternatives, Springer, New York, 1995.
[18] S. Wang, Theory of Linear Models and its Applications, Anhui Education Press, China, 1987 (in Chinese).
[19] S. Wang, Z. Jia, Inequalities in Matrix Theory, Anhui Education Press, China, 1994 (in Chinese).